

DOI: [10.14515/monitoring.2020.6.976](https://doi.org/10.14515/monitoring.2020.6.976)



В. Б. Звоновский, М. В. Григорьева

ПРАКТИКА ПРОВЕДЕНИЯ РОБОТИЗИРОВАННЫХ РЕПРЕЗЕНТАТИВНЫХ ТЕЛЕФОННЫХ ОПРОСОВ

Правильная ссылка на статью:

Звоновский В. Б., Григорьева М. В. Практика проведения роботизированных репрезентативных телефонных опросов // Мониторинг общественного мнения: экономические и социальные перемены. 2020. № 6. С. 73—96. <https://doi.org/10.14515/monitoring.2020.6.976>.

For citation:

Zvonovsky V. B., Grigorieva M. V. (2020) Experiences With IVR Representative Surveys. *Monitoring of Public Opinion: Economic and Social Changes*. No. 6. P. 73–96. <https://doi.org/10.14515/monitoring.2020.6.976>. (In Russ.)

ПРАКТИКА ПРОВЕДЕНИЯ РОБОТИЗИРОВАННЫХ РЕПРЕЗЕНТАТИВНЫХ ТЕЛЕФОННЫХ ОПРОСОВ

ЗВОНОВСКИЙ Владимир Борисович — доктор социологических наук, профессор, заведующий кафедрой социологии и психологии, Самарский государственный экономический университет, Самара, Россия; президент, Фонд социальных исследований, Самара, Россия

E-MAIL: zvb@socio-fond.com

<https://orcid.org/0000-0002-8805-9028>

ГРИГОРЬЕВА Мария Владимировна — кандидат политических наук, доцент кафедры социологии и психологии, Самарский государственный экономический университет, Москва, Россия

E-MAIL: mariagrigorieva9@gmail.com

<https://orcid.org/0000-0003-2819-9368>

Аннотация. В статье представлен обзор современных зарубежных и отечественных практик проведения роботизированных телефонных опросов (IVR/T-ACASI) в социологических исследованиях, их длительности, выбора скорости и тембра речи, материального и нематериального стимулирования респондентов, формулирования приветствия, использования сенситивных тем и вопросов. На основе результатов четырех волн экспериментального опроса избирателей Санкт-Петербурга накануне выборов губернатора города ($N = 136, 387, 1256, 1529$), проведенного с июля по сентябрь 2019 г., в статье сделан вывод о применимости зарубежных исследовательских рекомендаций к российским реалиям, а также выявлены особенности типичного поведения отечественных

EXPERIENCES WITH IVR REPRESENTATIVE SURVEYS

Vladimir B. ZVONOVSKY^{1,2} — Dr.Sci. (Soc.), Professor, Head of Department of Sociology and Psychology; President

E-MAIL: zvb@socio-fond.com

<https://orcid.org/0000-0002-8805-9028>

Maria V. GRIGORIEVA¹ — Cand. Sci. (Pol.), Associate Professor at the Department of Sociology and Psychology

E-MAIL: mariagrigorieva9@gmail.com

<https://orcid.org/0000-0003-2819-9368>

¹ Samara State University of Economics, Samara, Russia

² Social Research Foundation, Samara, Russia

Abstract. The article overviews modern foreign and domestic practices in conducting automated telephone surveys (IVR/T-ACASI) as part of sociological research highlighting aspects such as duration, speaking rate and pitch, financial and non-financial incentives for respondents, introduction wording and use of sensitive questions. The article is based on the results of four waves of an experimental survey among the residents of St Petersburg ($N = 136, 387, 1256, 1529$) on the eve of the St Petersburg gubernatorial election, from July to September 2019. The authors argue that findings of the foreign studies can be applied to the Russian realities and detect specific features in the behavior of the Russian respondents. In particular, the article justifies that surveys conducted over a short time frame ensure the

респондентов. В частности, обосновано проведение опросов в конкретный узкий промежуток времени с целью обеспечения наиболее высоких показателей кооперации респондентов; продемонстрирована важность формулировок приветствия, позволяющих минимизировать отказы от интервью. В заключении проанализированы возможности использования данных, собранных в ходе роботизированного опроса для формирования электорального прогноза. Статья предлагает рекомендации по использованию технологии IVR в российских исследованиях и намечает направления для дальнейшего усовершенствования данного метода.

Ключевые слова: телефонный опрос, роботизированный опрос, технология IVR, телефонный опрос избирателей, методика IVR-опроса, доступность респондентов

highest possible rates of respondent cooperation. Besides that, using a proper survey introduction wording can help keep non-responses to a minimum. In conclusion, the authors describe how the data collected via automated surveys can be used in election forecasting. The article offers recommendations for using the IVR technique in the Russian studies and outlines how this method can be further improved.

Keywords: telephone survey, automated survey, IVR, election survey, IVR survey method, respondent reach

Введение

Роботизированные телефонные опросы (или IVR¹-опросы) — один из новых и быстро набирающих популярность методов проведения социальных и маркетинговых исследований. Это комбинация технологий сбора данных, когда используется записанный или сгенерированный голос «интервьюера», а респондент отвечает путем нажатия клавиш телефона в тональном режиме или дает вербальные ответы, а система распознает и фиксирует их в базе данных.

Сегодня телефонные опросы — один из самых быстрых и дешевых способов первичного сбора данных в социологических, маркетинговых и электоральных исследованиях; они в наибольшей степени соответствуют требованиям к репрезентативным опросам. Именно в телефонных опросах, проводящихся в специально оборудованном помещении, где работу интервьюеров легко контролировать, а отклонение от инструкций можно оперативно обнаружить и предотвратить, реже встречаются фальсификации данных и ошибки измерения, связанные с «эффектами интервьюера» [Groves, 2004; Menold, Kemper 2013].

¹ IVR (interactive voice response (англ.) — интерактивный голосовой ответ) — это компьютеризированная технология сбора данных, также известная под аббревиатурой T-ACASI (telephone audio computer-assisted self-interviewing), но такое обозначение в литературе встречается гораздо реже.

Роботизированные телефонные опросы позволяют усилить эти преимущества: увеличить стандартизованность проведения интервью, существенно сократить стоимость исследования за счет уменьшения расходов на интервьюеров, радикально увеличить скорость сбора данных и охват исследуемой совокупности.

Еще недавно метод IVR как способ сбора данных проигрывал наиболее распространенному методу CATI (где вопросы респонденту задаются «настоящим» интервьюером) по показателям коэффициентов ответов и завершенных интервью [Curtin, Presser, Singer, 2005], но эти потери сегодня компенсируются широтой охвата (количеством телефонных номеров для дозвона) и кратчайшими сроками выполнения всех намеченных звонков, недоступных при использовании метода CATI [Kreuter, Presser, Tourangeau, 2008; Lee et al., 2018]. Столь же внушительно и сокращение расходов на сбор данных, так как отпадает необходимость нанимать интервьюеров. Все это — основные и несомненные технические преимущества метода IVR.

Но есть у него и другое крайне важное достоинство: IVR-опросы позволяют существенно снизить ошибки измерения, связанные не только с «эффектами интервьюера», но и с «эффектами респондента». Одним из основных источников такого рода ошибок являются ситуации, когда участников опроса просят отвечать на вопросы, затрагивающие так называемые чувствительные темы². Общаясь с живым человеком, респондент может побояться или просто не захотеть давать честный ответ, либо будет стремиться дать ответ, по его мнению, наиболее желательный, социально одобряемый. А роботизированные телефонные опросы обеспечивают полную конфиденциальность. Метод IVR может стать незаменимым, если в исследование включены вопросы, варианты ответов на которые содержат стигматизирующие в обществе характеристики, — например, затрагивающие тему здоровья и предполагающие наличие «неудобных» заболеваний [Shaw, Verma, 2007].

Отвечая на «чувствительные вопросы» роботу, респонденты не только с большей вероятностью сознаются в чем-то, что воспринимается как социально неодобряемое, но и, не испытывая потребности производить хорошее впечатление на интервьюера, менее склонны приписывать себе социально одобряемые, желательные характеристики и завышать свои социальные достижения [Kreuter, Presser, Tourangeau, 2008]. Поэтому некоторые исследователи предлагают сообщать респонденту, что разговор будет идти с роботом, в самом начале интервью и подчеркивать анонимность такого разговора [Amaya et al., 2018].

² Чувствительные, или чувствительные, темы (sensitive topics) — собирательный термин, касающийся вопросов, которые могут вызвать у респондентов негативные эмоции, нежелание давать честные ответы, а то и вовсе участвовать в опросе. Термин не получил однозначного определения в научной литературе. Как правило, принадлежность того или иного вопроса к числу чувствительных определяется по трем различным, но взаимосвязанным аспектам: 1) назойливость — респонденты могут счесть вопрос оскорбительным или просто недопустимым со стороны той или иной организации (например, вопрос о религиозных убеждениях респондента); 2) опасность — респонденты могут опасаться, что предоставляемая ими информация попадет в нежелательные руки (например, информация об уровне и источниках дохода окажется в распоряжении налоговых служб или агентов рынка; или информация о вредных привычках и сексуальной жизни подростка поступит администрации его школы или будет подслушана его родителями в ходе интервью); 3) стигматизация — вопросы могут потребовать от респондента признания в социально неодобряемых характеристиках (признание в наличии определенных заболеваний или согласие с непопулярным, порицаемым в обществе мнением). Отсутствие четкого определения термина «чувствительные темы» связано также с тем, что набор таких тем будет различным для разных стран и даже для разных социальных групп населения [Lavrakas, 2008].

IVR в сравнении с разговором с живым человеком предпочитают люди с низким уровнем доходов (это единственное различие между группами респондентов, выделенными по стандартным социо-демографическим показателям, в отношении степени их готовности проходить IVR-опрос). Других отличий в качестве получаемой методом IVR информации по сравнению с данными, получаемыми в ходе традиционных телефонных интервью, выявлено не было [Pariyo et al., 2019].

Методические особенности IVR интервью

Перед исследователями, проводящими роботизированные опросы, стоит ряд специфических задач, которым необходимо уделять внимание, чтобы использовать преимущества метода в полной мере. К ним относятся:

- качество технологического и программного обеспечения процесса проведения опроса данного типа;
- особые требования к дизайну опросника;
- необходимость учета психологических особенностей респондентов для стимулирования их к сотрудничеству и удержания их внимания до конца интервью.

Поскольку при опросе методом IVR отсутствует живой интервьюер, в чьи задачи среди прочего входят налаживание кооперации, разъяснение сути вопросов, удержание внимания и интереса респондента, то все эти моменты должны быть заранее продуманы и решены в составленном и записанном опроснике [Рогозин, Вырская, 2019].

Наиболее значимую роль в IVR-интервью играет содержание приветствия, поскольку именно от него во многом зависит решение респондента об участии в опросе.

Записанный голос должен вызывать доверие, а формулировки вопросов и предложений должны свидетельствовать о компетентности людей, проводящих опрос. При этом необходимо избегать ситуации, в которой респондент может почувствовать себя в роли подчиненного. **В обращении к респонденту необходимо избегать языка субординации, следует использовать стиль общения между равными взрослыми людьми** (Adult-to-Adult Communication Style). Не стоит использовать формулировки, которые потенциальные респонденты могут счесть унижительными, например: «Вы должны пройти этот опрос сегодня!» или «Чтобы у нас была возможность помочь решить проблемы школьного образования в вашем городе, вам необходимо пройти этот опрос». Правильный вариант звучал бы так: «Ваши ответы смогут помочь лучше понять проблемы, стоящие перед школами в нашем городе» [Comley, 2006].

Необходимо внимательно относиться к тому, что **цели, с которыми проводится опрос, могут быть понятны и близки одним людям, но оставлять равнодушными других и даже вызывать у них отторжение.** Поэтому в приветствии необходимо подчеркнуть положительные для потенциальных респондентов моменты, а отрицательные сделать менее заметными. Эти моменты в приветствии должны корректироваться по ходу проведения исследований и наблюдения за реакциями участников опроса, чтобы в дальнейшем повышать вероятность получения ответов и завершенных интервью [Groves, Singer, Corning, 2000].

Полезным может оказаться **упоминание в приветствии названия организации, которая проводит или спонсирует проведение опроса**. Как правило, население большинства стран считает заслуживающими доверия и имеющими право задавать вопросы различные государственные структуры, а также крупные и хорошо известные местные университеты. Впрочем, спонсорство любой — некоммерческой или коммерческой, правительственной или общественной — организации стоит обозначить в приветствии, поскольку повышает доверие респондента к опросу и увеличивает шанс на его кооперацию [Dillman, Smyth, Christian, 2014]. Положительно на ответах респондентов может сказаться предоставление более подробных сведений о целях опроса, однако не рекомендуется включать такую информацию в сам опрос. Например, не стоит предусматривать возможность прослушать дополнительную информацию об опросе путем нажатия той или иной клавиши телефона прямо во время проведения IVR-интервью. Лучше ограничиться указанием на ресурсы, где эта информация может быть получена.

Если у конкретного респондента уже есть определенная история участия в опросах, то **указание на сходство данного опроса с предыдущими или на связь респондента с организацией, проводящей текущий опрос, положительно скажется на готовности респондента принять участие** и в этот раз [Lynn, 2009].

Еще один способ повысить кооперацию респондентов — **денежное или иное материальное поощрение** участия. Зарубежные исследователи указывают, что лучше предоставлять такие стимулы заранее (предваряющая дозвон почтовая или смс-рассылка) и «безвозмездно», чтобы они мотивировали принять участие, а не были приманкой, вынуждающей пройти опрос. Ученые подчеркивают, что таким способом доверие респондента завоевывается более эффективно и, следовательно, повышается его готовность к сотрудничеству. Предоставляя стимул вместе с запросом до прохождения опроса, исследователь как бы говорит, что это всего лишь «небольшой знак признательности», и таким образом демонстрирует доверие и уважение к потенциальным респондентам, которые имеют возможность просто положить деньги в карман, уклонившись от участия [Church, 1993; Singer, Ye, 2013]. Что касается размера денежного поощрения, то, как указывают зарубежные исследователи, два доллара будут лучше, чем один, но вот сумма в пятнадцать долларов за десятиминутный опрос скорее вызовет подозрения, нежели доверие, и с большей вероятностью станет удерживающим от кооперации фактором [Trussell, Lavrakas, 2004]. Очевидно, что такой стимул не будет адекватным для массовых опросов, поскольку трудно подобрать размер вознаграждения, в одинаковой степени стимулирующий группы с разным уровнем доходов.

Готовность респондентов проходить телефонные опросы значительно уменьшилась со временем, о чем свидетельствуют неуклонно снижающиеся показатели ответов [Curtin, Presser, Singer, 2005]. В первую очередь это заметно в наиболее молодой возрастной группе, представители которой больше привыкли иметь дело с текстовыми сообщениями, нежели с голосовыми вызовами³. Исследователям стоит задуматься о предоставлении респонденту возможности пройти опрос

³ Lenhart A., Ling R., Campbell S., Purcell K. Teens and Mobile Phones: Text Messaging Explodes as Teens Embrace It as the Centerpiece of Their Communication Strategies With Friends // Pew Research Center. 2010. April 20. URL: <https://www.pewresearch.org/internet/2010/04/20/teens-and-mobile-phones/>.

в удобном для него режиме [Olson, Smyth, Wood, 2012]. Это может быть использование тонального набора, позволяющего выбирать варианты ответов путем нажатия клавиш, либо других возможностей смартфонов (но тогда остро встает вопрос разработки дизайна анкет, заполнение которых стало бы увлекательным для респондента) [Link et al., 2014]. Тем не менее часть исследователей полагают, что если предоставить респонденту возможность выбрать режим прохождения опроса, это скорее снизит коэффициент ответов, так как необходимость переключаться с одного способа на другой вызовет желание просто прервать интервью в момент переключения [Millar, Dillman, 2011].

Представления о допустимой длине анкеты сильно разнятся в зависимости от характера исследования. Иногда она доходит до 40 вопросов, но в таком случае участие в опросе напрямую связано с интересами респондента, который имеет отношение к исследуемой организации. К примеру, это могут быть опросы о состоянии здоровья, проводимые клиникой, в которой наблюдается респондент, или опросы выпускников конкретного университета [Kreuter, Presser, Tourangeau, 2008; Shaw, Verma, 2007]. В остальных же случаях рекомендуется сократить длину опросника, задавая только необходимые для конкретного исследования вопросы и добиваясь ответов на минимально необходимом уровне детализации. Исследователь не должен пытаться утаить от респондента предполагаемые временные затраты, поскольку это, согласно собранным данным, не влияет положительно на коэффициент завершенных интервью [Dillman, 2011; Pariyo et al., 2019]. Время, необходимое для прохождения опроса, — ценный ресурс для респондента, который готовы потратить далеко не все. Следует учитывать, что нежелание отвечать усиливается, если сами вопросы сложны для понимания. Стремление получить ответы на подробные и разноплановые вопросы, собрать всю возможную информацию «про запас» входит в конфликт с лимитом терпения респондентов. Сегодня повседневное общение стало намного короче и быстрее, для него требуется немногим более щелчка мыши или нажатия клавиши на экране смартфона. Поэтому допустимы интервью длиной в 10—12 вопросов (и все они должны напрямую соотноситься с целями, ради которых и был организован опрос). Сами вопросы должны предполагать ответы «да/нет», выбор из ограниченного числа вариантов или выбор наиболее подходящей оценки из предложенных вариантов. Стоит избегать формулировок с ранжированием, так как респонденту будет сложно удерживать варианты в памяти. Лучше не злоупотреблять порядковыми шкалами, ибо люди склонны запоминать и чаще выбирать крайние категории [Dillman et al., 2009]. Предпочтительны повышенная скорость речи и сокращение пауз между вопросами, чтобы не истощить терпение участников опроса [Amaуa et al., 2018; Charoenruk, Olson, 2018]⁴.

В российской практике использование роботов в массовых телефонных опросах носит единичный и очень узкий характер. Так, Д. Рогозин и М. Вырская рассказывают о своей попытке с помощью роботизированного опроса рекрутировать выборку респондентов, которые впоследствии должны были быть опрошены ме-

⁴ См. также: Lenhart A., Ling R., Campbell S., Purcell K. Teens and Mobile Phones: Text Messaging Explodes as Teens Embrace It as the Centerpiece of Their Communication Strategies With Friends // Pew Research Center. 2010. April 20. URL: <https://www.pewresearch.org/internet/2010/04/20/teens-and-mobile-phones/>.

тодом обычного телефонного опроса [Рогозин, Вырская, 2019]. Методологическое описание репрезентативных телефонных опросов с использованием IVR по состоянию на август 2020 г. не обнаружено.

Итоги обзора достигнутых в разных странах результатов проведения роботизированных опросов можно представить следующим образом (см. табл. 1).

Таблица 1. *Достоинства и недостатки роботизированного интервью (IVR)*

Достоинства	Недостатки	Аспекты, требующие внимания
Высокая стандартизированность	Небольшая длительность интервью	Приветствие, включая обозначение целей проведения опроса
Низкая стоимость	Доминирование дихотомических шкал	Скорость, тембр голоса
Высокая скорость		Обозначение спонсора опроса
Низкая чувствительность		Материальное вознаграждение

Как видим, снижение стоимости и времени на проведение сбора данных с помощью IVR заставляет нас использовать интервью меньшей длительности по сравнению с традиционными телефонными интервью и ограничивать себя простейшими дихотомическими шкалами. С другой стороны, высокая стандартизация IVR-опроса не исключает формирования искусственного разнообразия голосов и тембра синтезированного голоса.

Тестирование метода IVR интервью на выборах губернатора Санкт-Петербурга в 2019 г.

Методология исследования

Для проведения тестовых измерений была выбрана генеральная совокупность избирателей Санкт-Петербурга, что обусловлено рядом причин, к ключевым из которых относятся:

- ее большой объем и разнообразие единиц, позволяющие генерализировать результаты теста на население большей части страны;
- в этой совокупности на регулярной основе проводились волны опроса методом традиционного CATI, что дало возможность сопоставить результаты двух массивов данных, собранных на одной популяции и с одним основанием выборки (в обоих случаях использовалось одно и то же основание для производства выборки телефонных номеров), но с применением двух методов;
- Санкт-Петербург традиционно относится к «трудным» для проведения опросов регионам России по причине высокой мобильности горожан, их включенности в разнообразные городские практики. Это приводит к низкой кооперации и большой доле отказов. Такие особенности населения города позволяют предположить, что достигнутые на этой популяции результаты тестирования роботизированного

опроса не будут значительно отличаться от результатов аналогичного тестирования по всей стране.

При всех особенностях избирателей Санкт-Петербурга как генеральной совокупности, результаты, полученные на этой популяции, можно распространить на другие крупные российские города в силу сходства практик коммуникативного поведения населения, в частности общения посредством телефонной связи.

Цели, задачи и методические предпосылки исследования

Целью проведенного исследования было тестирование роботизированной системы сбора первичных данных в ходе массового социологического опроса. Для этого был проведен роботизированный опрос населения крупного российского города одновременно с опросом, реализованным методом традиционного CATI-опроса.

Были поставлены следующие задачи исследования:

- построить оптимальный дизайн анкеты;
- сформировать адекватный набор используемых шкал;
- определить оптимальное время проведения опроса;
- сравнить результаты традиционного телефонного и роботизированного опросов.

Объект исследования — массивы первичных данных, собранных роботом и интервьюерами. **Предмет исследования** — расхождения между данными, собранными с помощью роботов и с помощью интервьюеров, и факторы, влияющие на степень обнаруженных расхождений.

В данной статье под роботом мы будем понимать программный комплекс, предназначенный для синтеза и распознавания человеческой речи, а также сигналов телекоммуникационных сетей. Распознанные ответы респондентов составляют анализируемый массив данных, а зафиксированные и проинтерпретированные сигналы телекоммуникационных сетей формируют массив параданных, дающий возможность оценить качество проведения самого опроса: отбора и достижимости различных групп респондентов (по ключевым социально-демографическим характеристикам).

В зарубежных социологических исследованиях технические аспекты, связанные с настройкой программного обеспечения, как правило, игнорируются. Исключение составляют медицинские исследования, сосредоточенные на решении сугубо медицинских задач (есть ли у пациента те или иные симптомы заболевания, какие он принимает лекарства и иная информация, позволяющая в автоматическом режиме сформировать предварительный диагноз), и в таких проектах респонденты, как правило, напрямую заинтересованы в сотрудничестве [Falconi, Johnston, Hogg, 2016; James et al., 2018]. Но и из этих работ можно получить ценные советы. Рекомендуется в ходе первоначального тестирования системы привлекать координаторов, которые будут подключаться к интервью и отслеживать правильность распознавания роботами ответов, поскольку настройка распознавания речи требует тщательного рассмотрения нескольких переменных. К ним относится установка порога вероятности, необходимая для идентификации ответа, а также конкретные принятые ответы (то есть «да», «конечно», «хорошо»).

Все системы распознавания речи используют вероятностные модели, такие как Hidden Markov model⁵.

Это означает, что робот никогда не может быть на 100 % уверен в том, что данное слово было произнесено. Оценивается только вероятность этого. Например, респондент говорит «выборы». Робот сравнивает произношение с имеющимися у него стандартами. Однако совпадение фонем не может быть стопроцентным, поэтому, давая задание роботу, необходимо принять решение о том, какой уровень вероятности допустим, чтобы распознавание состоялось. Низкие настройки (ответ фиксируется даже когда установленная компьютером вероятность произнесения именно этого слова невысока) несут в себе риск неправильного распознавания слов, а высокие настройки (вероятность произнесения конкретного ответа стремится к 100 %) могут привести к тому, что правильное слово не будет распознано и зафиксировано. Это испытание методом проб и ошибок должно включать респондентов обоего пола, носителей различных акцентов и диалектов, которые предположительно могут оказаться в выборке [Willig et al., 2013].

Другой важный момент — возможность прервать робота и дать ответ, как только респондент услышит подходящий вариант, а не прослушивать все пункты до конца. Установление того, какие части сценария IVR допустимо «прерывать», а какие нет (когда необходима уверенность, что человек прослушал вопрос или фразу целиком, например при выполнении условия информированного согласия), сократит временную нагрузку на участников и повысит полноту и точность информации [ibidem].

Положительным фактором является использование call-центра для исходящих вызовов локальных телефонных префиксов, которые вызывают большее доверие у респондентов. В частности, известно, что звонки с неизвестного «московского» номера могут просто игнорироваться жителями Санкт-Петербурга.

В рамках проведенного нами эксперимента было принято несколько решений относительно выбора тех или вариантов работы робота.

Решение не оповещать участников опроса о том, что их опрашивает робот, оказалось не очень удачным. С одной стороны, не менее 20 % респондентов отвечали на вопросы до конца, будучи уверены в том, что общаются с живым человеком. Другая их часть (не более 5 %) быстро опознавали робота, приветствуя его словами «Здравствуй, Алиса!» или «ОК, Google».

Работа была организована таким образом, чтобы полевая часть исследования заняла как можно меньше времени. Это исключало возможность второй и последующих попыток контакта. Участники опроса не получали какого-либо вознаграждения за участие в опросе. Также им не анонсировали предстоящий звонок из опросного центра. Мы никак не стимулировали людей к участию в опросе, кроме как приветствием, чтобы иметь возможность оценить, приветствия какого рода наиболее эффективны в плане побуждения респондентов к участию в опросе.

С первой волны и во всех последующих волнах исследования мы использовали только характерные для Санкт-Петербурга телефонные префиксы.

⁵ Hidden Markov model (скрытые марковские процессы) представляет собой сравнения некоторых известных стандартов (в случае речи — частотной характеристики звучания различных фонем) и реальных фонем, произносимых (в случае опроса) респондентами.

У респондентов появлялась возможность прервать работу робота в случае, когда варианты ответов были очевидны и не требовали прослушивания всей шкалы ответов (например, пола, возраста и района проживания), а также при повторном зачитывании вопроса и вариантов ответа, если первый ответ был не получен или не распознан.

Особенности формирования выборки

Описываемое исследование было проведено в четыре волны. Генерация телефонных номеров на всех этапах производилась ВЦИОМ. Массив номеров включал в себя операторов мобильной и стационарной связи с кодами (DEF- и ABC-соответственно), приписанными к Санкт-Петербургу и Ленинградской области. Соотношение номеров стационарной и мобильной связи, а также время звонков во всех волнах представлено в таблице 2.

Таблица 2. Сроки и объемы стартовой выборки номеров проведенных волн измерения

	18—21 июня Вт-Чт	24—26 августа Пт-Вс	2 сентября Пн	8 сентября Вс
Время звонков	17.00—20.00	17.40—22.00	15.00—19.00	16.00—21.00
Предоставлено номеров всего. Из них:	50 544	399 342	496 003	598 615
<i>Мобильных (второй уровень)</i>	33 042	399 342	490 000	448 610
<i>Стационарных (второй уровень)</i>	17 502	—	6 003	150 005

Как показано в таблице 3, в первой волне робот собрал 136 законченных интервью на основе выданных ему 50 544 номеров, что составило 0,27 %. В третьей было получено уже 1 256 законченных интервью из 496 003 номеров стартовой выборки (0,25 %). С четвертой волной такое сравнение не релевантно, поскольку в ней полное интервью проходили только те, кто участвовал в выборах губернатора Санкт-Петербурга. С точки зрения доли законченных интервью от числа всех предоставленных номеров значимой динамики по итогам трех волн исследования не обнаружено, а вот доля законченных интервью среди тех, кому было зачитано приветствие (response rate), увеличилась с 2,7 % до 8,6 %. То есть кооперация робота с респондентами довольно быстро заметно выросла. Аналогичный показатель для интервьюеров в этот период колебался между 8,2 % и 11,6 %. Как видим, в сотрудничестве с респондентами роботы уже сегодня показывают значения, близкие к традиционным опросам. Даже небольшой по длительности период обучения роботов дает очень хороший результат. Это позволяет предполагать, что дальнейшее совершенствование распознавания позволит добиться лучших показателей.

Таблица 3. Основные параданные волн исследования

Результаты работы робота	1 волна: 18—21 июня	2 волна: 24—26 августа	3 волна: 2 сентября	4 волна: 8 сентября
Полные интервью (есть ответы более чем на 50%)	136	387	1256	1529
Неполные интервью (нет ответов на 50% вопросов)	204	301	1546	3034
Прерванные интервью (ответы только на 2 первых вопроса)	—	158	728	1805
Номеров выдано	50 544	399 342	496 003	598 615
Доля полных интервью от числа номеров	0,27%	0,10%	0,25%	0,26%
Доля полных интервью от числа состоявшихся контактов	2,7%	3,8%	8,6%	2,9%
Результаты работы традиционного САП	1 волна: 24—28 июня	2 волна: 26—27 августа	3 волна: 1 сентября	4 волна: 6 сентября
Полные интервью	1 209	1 209	1 209	1 209
Номеров выдано	189 542	982 692	996 423	1 024 303
Доля полных интервью от числа номеров	0,63%	0,12%	0,12%	0,12%
Доля полных интервью от числа состояв- шихся контактов	8,2%	11,8%	11,6%	9,2%

Кооперация респондентов по часам

Относительно времени проведения телефонных опросов можно сделать следующие предположения:

- респонденты неохотно будут отвечать, если находятся на рабочем месте или за рулем (но, вероятно, готовы отвечать на вопросы, если являются пассажирами);
- респонденты охотнее будут отвечать в выходные дни и в вечерние часы. Однако, вероятно, после 20.00 готовность респондентов к участию в опросе будет снижаться.

На основании этих предположений были сформулированы две гипотезы:

1. Доля прерванных интервью уменьшается по мере окончания рабочего дня и в часы, которые у большинства жителей города тратятся на дорогу с места работы домой, то есть примерно до 19.00 для такого большого города, как Санкт-Петербург. Доля отказов — вербальных и невербальных — также уменьшается.

2. Доля не взявших трубку будет увеличиваться после 19.00.

Как показано в таблице 4, доля прерванных интервью в первой волне снизилась с 12,5% в период с 17.00—18.00 до 8,0% в период 18.00—19.00. Однако в другие волны подобной динамики не наблюдалось. Возможно, по той причине, что рабочими днями были лишь последний день второй волны (28 августа) и третья волна (2 сентября). При этом в третьей волне в эти часы существенно сократилась доля вербальных отказов от интервью (с 60% в период 16.00—17.00 до 39%-37%

в следующие два часа — 17.00—18.00 и 18.00—19.00). Таким образом, гипотезы в целом подтвердились. Респонденты отказываются от интервью в рабочее время и когда едут с работы домой. Соответственно, для проведения телефонных опросов, как роботизированных, так и традиционных, у исследователей есть небольшой промежуток времени между 19.00 и 21.00 местного времени, когда респонденты демонстрируют наибольшую готовность к сотрудничеству.

Таблица 4. Динамика количества прерванных интервью и отказов по часовым интервалам

	16.00— 17.00	17.00— 18.00	18.00— 19.00	19.00— 20.00	20.00— 21.00	21.00— 22.00
<i>Доля прерванных интервью (данные собирались по часам в первой и четвертой волнах), в %</i>						
1 волна	—	12,5	8,0	5,9	—	—
4 волна	1,6	1,8	1,8	1,8	1,8	—
<i>Доля вербальных отказов, в %</i>						
2 волна	—	54	48	47	55	67
3 волна	60	39	37	—	—	—
4 волна	51	49	51	45	52	—
<i>Не взяли трубку среди всех набранных номеров, в %</i>						
1 волна	—	69	66	56	—	—
2 волна	—	45	41	46	51	61
3 волна	65	64	55	—	—	—
4 волна	63	62	63	53	64	—

Относительно доли не взявших трубку отметим, что в первой волне она сократилась с 66% в 18.00—19.00 до 56% в 19.00—20.00, что может быть и результатом возросшей кооперации после возвращения респондентов домой (работающее население составило от 42% до 54% выборочной совокупности в различных волнах). Аналогичную динамику мы имеем в третьей волне, где этот показатель уменьшился с 64% в 17.00—18.00 до 55% в 18.00—19.00. Во второй волне доля вербального отказа не сокращалась в эти часы (16.00—19.00), возможно, по причине того, что опрос проходил с пятницы по воскресенье, и рабочих часов было немного. Однако после 19.00 доля отказов стала расти с 46% в 19.00—20.00 до 61% в 21.00—22.00. Столь же быстро в первой волне увеличивалась доля вербальных отказов — с 47% до 67%. В четвертой волне, которая проводилась в воскресенье, доля не взявших трубку оставалась постоянной, за исключением часа между 19.00 и 20.00, причину чего объяснить сложно. Таким образом, можно с большой долей уверенности утверждать, что кооперация респондентов возрастает по мере окончания рабочего дня и возвращения с работы домой, но сокращается, если приглашение к опросу происходит после 19.00.

Как видим, сформулированные гипотезы подтверждаются: существует наиболее благоприятный промежуток для проведения телефонных опросов — между

18 и 19 часами в будние дни, но определение оптимального времени для максимальной кооперации требует более точных измерений и, вероятно, зависит от времени года. Очевидно, что роботизированные опросы, не чувствительные к числу операторов, задействованных в единицу времени, имеют однозначное преимущество, позволяя охватить все телефонные номера выборки в краткий промежуток времени, что невозможно для живых интервьюеров в опросах по методу CATI.

Привлечение респондентов в опрос (скринирование)

Хороший результат дают роботы для получения ответов на все вопросы, сформулированные в анкете. Как видно из таблицы 5, доля завершенных интервью колебалась между первой и третьей волнами от 36 % до 46 % от числа начатых интервью (был получен ответ хотя бы на первый вопрос), в то время как операторам удавалось довести интервью до конца с респондентами куда реже: их доля незаконченных интервью колебалась между 8 % и 9%. Конечно, их интервью было гораздо длиннее, но сам по себе показатель кооперации с роботом у современного респондента очень высокий, что позволяет рассчитывать на его дальнейший рост в будущем при совершенствовании алгоритмов.

Таблица 5. Ключевые параданные по роботизированному и традиционному опросам

Результаты работы робота	1 волна: 18—21 июня	2 волна: 24—26 августа	3 волна: 2 сентября	4 волна: 8 сентября
Полные интервью (есть ответы более чем на 50 %)	136	387	1256	1529
Неполные интервью (нет ответов на 50 % вопросов)	204	301	1546	3034
Прерванные интервью (ответы только на 2 первых вопроса)	-	158	728	1805
Не были отобраны по квоте по проживанию в Санкт-Петербурге	4*	533	1243	1071
Не были отобраны по возрасту (менее 18 лет)	40**	82	374	343**
Соединено (зачитано приветствие/приглашение)	4963	10287	14613	52987
Неверные номера (автоответчик либо ошибка вызова)	10556	22296	57844	52893
Использовано номеров	50294	90877	496003	332720
Результаты работы традиционного CATI	1 волна: 24—28 июня	2 волна: 26—27 августа	3 волна: 1 сентября	4 волна: 6 сентября
Полные интервью (есть ответы более чем на 50 %)	1209	1209	1209	1209
Соединено (зачитано приветствие/приглашение)	14713	10088	10348	13006

Результаты работы робота	1 волна: 18—21 июня	2 волна: 24—26 августа	3 волна: 2 сентября	4 волна: 8 сентября
Не были отобраны по квоте по проживанию в Санкт-Петербурге	664	419	532	1 040
Не были отобраны по возрасту (менее 18 лет)	54	18	17	24
Неверные номера (автоответчик либо ошибка вызова)	1 967	32 883	8 745	56 887
Предоставлено номеров	189 542	982 692	996 423	1 024 303

* Вопрос о месте проживания был задан третьим по счету, к этому времени часть респондентов прервала интервью раньше. Во второй и последующих волнах этот вопрос задавался первым.

** Был только один прямой вопрос о возрасте. Ошибки распознавания более вероятны, чем при ответе на дихотомический вопрос о том, исполнилось ли 18 лет.

Прерывание интервью

Важным показателем качества проведенной полевой работы является доля прерванных интервью, то есть интервью с респондентами, подошедшими по скринингу, начавшими отвечать на вопросы, но переставшими в какой-то момент давать ответы и/или бросившими трубку.

Интервью прерываются по разным причинам. Это и технические сложности (например, качество связи), и ошибки распознавания ответов роботом (как в случае неожиданных формулировок), и нежелание отвечать на вопросы по данной теме. Чаще всего мы имеем совместное действие всех трех причин. Например, акцент или дефект речи респондента в дополнение к помехам на линии сильно затрудняют распознавание ответов роботом. Влияние первых двух причин равномерно распределено по времени интервью и «осыпание» численности респондентов происходит монотонно от вопроса к вопросу. Таким образом, выборка «осыпается» под влиянием восприятия текста или темы вопроса, при котором интервью прерывается по инициативе респондента. Следует уточнить, что были случаи, когда связь обрывалась вследствие программной ошибки робота, но такие случаи также были равномерно распределены по всей длине интервью.

В отличие от последующих волн, в первой волне отсутствие ответа на первый вопрос приводило к прекращению интервью (см. рис. 1). Вероятно, по этой причине «осыпание» происходило с более высокой позиции (64%), хотя и примерно с такой же скоростью (угол наклона прямой к оси абсцисс). В четвертой волне целевая аудитория была более узкой (только участвовавшие в голосовании на выборах губернатора), поэтому уменьшение выборки было значительнее. Если же говорить о второй и третьей волнах, то «осыпание» происходило с уровня около 55% до примерно 35%. Иначе говоря, процесс скринирования приводил к потере примерно двух из каждых трех отозвавшихся респондентов. В какой-то мере это естественное «осыпание», поскольку часть отозвавшихся не подходит по условиям отбора (например, по возрасту либо по месту проживания). В ходе основного интервью потери должны быть меньше.

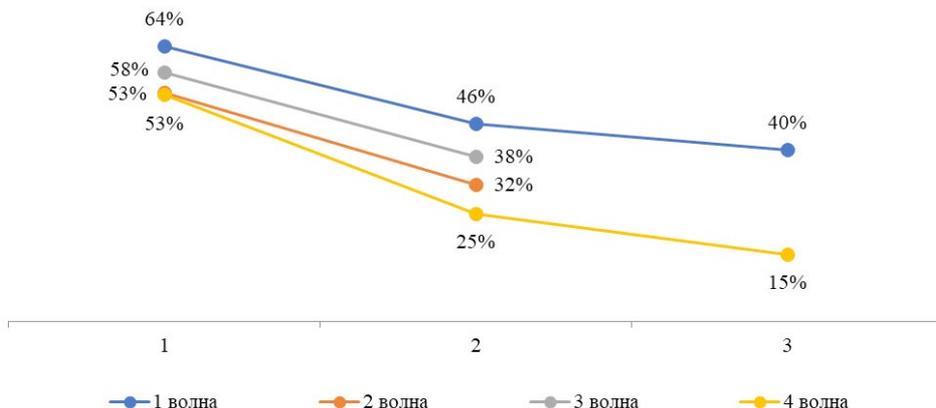


Рис. 1. «Осыпание» выборки (доля давших ответ от числа начавших отвечать на вопросы) на скрининговой части анкеты с увеличением номера вопроса (по оси абсцисс)

Как видно на рисунке 2, к концу интервью первой волны осталось 55 % от числа прошедших скрининговую часть, к концу второй, содержательной, части интервью осталось 58 %, третьей — 81 %, четвертой — 78 %. Во второй части интервью «осыпание» не только существенно слабее, но и с каждой волной все меньше. Можно предположить, что робот с каждой волной улучшал распознавание ответов, что позволяло все большему числу респондентов пройти интервью до конца. И дальнейшая работа робота, накопление базы ответов, будут улучшать качество его работы, а значит, увеличивать долю законченных интервью.

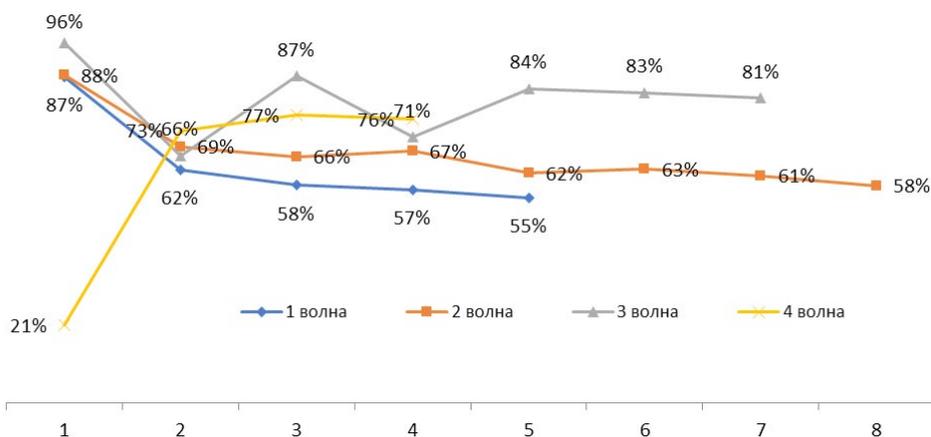


Рис. 2. «Осыпание» (доля давших ответ от числа начавших отвечать на вопросы) выборки на основной части анкеты с увеличением номера вопроса (по оси абсцисс)

Зарубежные исследователи пришли к выводу, что в роботизированных телефонных опросах прерывание интервью и сенситивность тем, затрагиваемых в вопросах, не связаны между собой, а само прерывание объясняется лишь усталостью респондента [Kreuter et al., 2008].

Наши данные показывают, что это не так. За исключением скрининговых вопросов, наибольшие потери в численности ответивших были понесены на третьем вопросе (о знании самого факта проведения выборов), когда еще рано говорить об усталости отвечающего, изначально готового пройти опрос. Но именно в третьем вопросе («Знаете ли Вы, когда состоятся следующие выборы главы Санкт-Петербурга?») впервые для респондентов звучала главная тема интервью — выборы губернатора Санкт-Петербурга. Она оказалась нежелательной для респондентов не столько из-за ее предполагаемой сенситивности, сколько по причине ее периферийности для участников опроса. Среди записанных ответов, подтверждающих это объяснение, встречались такие, как «а зачем мне это знать?», «мне это неинтересно» и т. п. Наблюдалась и более агрессивная реакция на данный вопрос. Возможно, «усталость» здесь присутствует как перенасыщенность этой темой в предвыборный период, но, скорее всего, мы имеем дело с теми, для кого политическая тема в принципе не актуальна⁶.

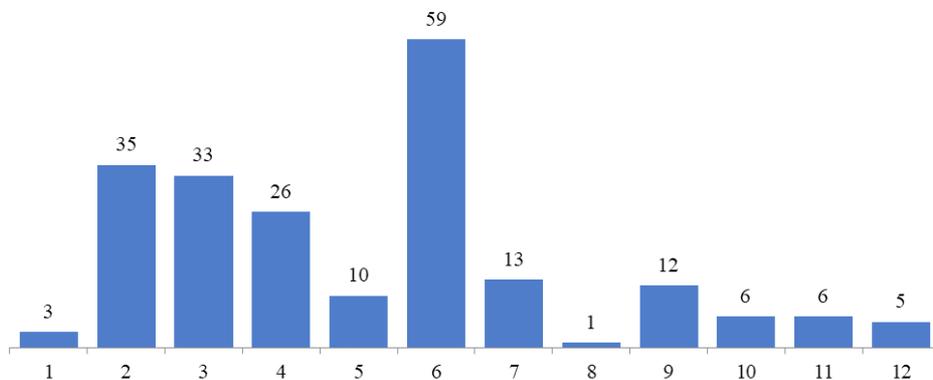


Рис. 3. «Осыпание» выборки (число прервавших интервью) в зависимости от номера вопроса в третьей волне опроса

Как видно из рисунка 3, наиболее критичным по числу респондентов, прервавших интервью, стал вопрос (чаще всего это был шестой по порядку вопрос) о том, за кого из кандидатов он намерен проголосовать на выборах. Это подтверждается репликами отвечающих, заявляющих, что они не будут отвечать на этот вопрос и что «это только их личное дело».

⁶ Поскольку в приветствии не говорилось, что опрос проводит робот, многие респонденты не понимали, что имеют дело с ним. О конфиденциальности ответов в приветствии речь также не шла, соответственно, не для всех участников опроса чувствительность темы была сглажена уверенностью в анонимности.

В дальнейшем число прерванных интервью резко сокращается. Казалось бы, именно после первого «электорального» вопроса уже можно говорить об усталости респондента, но ее влияние на решение продолжать или прервать интервью несравнимо с тем, какое влияние на решение респондентов прервать интервью оказывают сами формулировки сценария.

Можно предположить, что «осыпания» выборки также связаны с тем, как именно мы спрашиваем: предыдущие исследования показали, что для респондента утомительно выслушивать по телефону и запоминать многочисленные варианты предлагаемых ответов. А вопрос о предпочитаемом кандидате содержал четыре варианта, в которых робот зачитывал полные имена кандидатов и их должности.

Отчасти это предположение подтверждается тем, что в третьей волне в 13 интервью были прерваны на восьмом вопросе, еще 12 — на десятом (всего было прервано 545). Восьмой касался образования респондента, десятый — частоты и продолжительности использования им интернета, и каждый содержал по шесть вариантов предлагаемых ответов. При этом на девятом вопросе (о том, работает ли респондент), который предлагал ответить лишь «да» или «нет», прервалось всего одно интервью. Очевидно, что все три вопроса касались личной информации и респондент не мог не знать ответов на них.

Стоит отметить, что на предшествующем седьмом вопросе такого «осыпания» выборки не происходило. Кроме того, вопросы с третьего по шестой, которые также касались выборов и демонстрировали большое количество прерываний интервью, были простыми с точки зрения восприятия респондентом. Можно предположить, что на долю прерванных интервью влияет не столько усталость респондента, сколько тема опроса.

Влияние приветствия на участие в опросе

В ходе данного исследования мы тестировали ряд теорий, касающихся привлечения респондентов к участию в социологических опросах.

В частности, теория социального обмена (social exchange theory) предполагает, что люди с большей вероятностью ответят на чей-то запрос, если верят, что вознаграждение за отклик превысит затраты на его выполнение. В соответствии с этой теорией, большинство людей принимает решение об участии в опросе (или о продолжении интервью), опираясь на множество соображений, которые учитывают ожидаемые выгоды и затраты, и веру в то, что в долгосрочной перспективе первые превысят вторые [Comley, 2006; Stafford, 2008]. Сокращение каких-то неудобств, связанных с прохождением опроса, желательно, но без выигрыша от участия в нем не сыграет существенной роли: люди предпочитают действовать, когда в их субъективном исчислении польза перевешивает расходы [Singer, 2011].

Некоторые теории также подразумевают, что люди действуют, в частности, из-за того, что получают удовлетворение, помогая другим, особенно, если бенефициарами выступают отдельные лица или группы (школа, сообщество, страна), к которым респондент испытывает сочувствие или симпатию; он может принадлежать к ним сам или получать от них долгосрочные выгоды. Оказывая помощь, люди также стремятся к реализации личных и, возможно, эгоистических интересов [Dillman, Smyth, Christian, 2014; Groves, Singer, Corning, 2000].

В соответствии с приведенными теориями мы составили три формы приветствия, которые тестировались в ходе исследования:

Приветствие № 1 — самое короткое, содержащее лишь представление проводящей опрос организации и просьбу принять участие в исследовании: «Добрый день. Центр изучения общественного мнения проводит опрос жителей Санкт-Петербурга по вопросам жизни города. Пожалуйста, уделите пару минут вашего времени на прохождение опроса». Случаев с таким приветствием насчитывается 292, из них 132 интервью (45%) не были доведены до конца, а были прерваны респондентами на том или ином этапе разговора.

Приветствие № 2 содержало прямую просьбу помочь протестировать систему опросов: «Добрый день. Центр изучения общественного мнения проводит опрос жителей Санкт-Петербурга по вопросам жизни города. Мы разработали новую систему опросов населения, и хотим, чтобы Вы помогли нам ее протестировать. Пожалуйста, уделите пару минут вашего времени на прохождение опроса». Случаев с таким вариантом приветствия в нашей выборке 278. Из них 113 интервью (41%) были прерваны респондентом в ходе задавания вопросов.

Приветствие № 3 подчеркивало факт проведения опроса с целью лучшего понимания интересов жителей города, что подразумевало долгосрочную выгоду для респондента: «Добрый день. Центр изучения общественного мнения проводит опрос жителей Санкт-Петербурга. Опрос займет пару минут, но собранные данные позволят нам лучше понимать интересы различных групп жителей города». С таким вариантом приветствия у нас 276 случаев, из которых респондентом были прерваны до окончания лишь 100 интервью (36%), что представляет собой самый низкий показатель прерванных интервью среди трех групп вариантов приветствия.

Таблица 6. Влияние приветствия на кооперацию респондента по волнам исследования

	Прерванное интервью				Завершенное интервью*			
	1 волна	2 волна	3 волна	4 волна	1 волна	2 волна	3 волна	4 волна
Приветствие 1 (краткое)	98	132	586	1109	27	130	426	506
Приветствие 2 (помощь)	169	113	471	983	52	128	421	508
Приветствие 3 (интерес)	161	100	535	983	52	128	409	515
ВСЕГО	428	345	473	1048	57	129	421	1529
Хи-квадрат, актуальное	29,81	9,27	15,23	5,97				

* Завершенным здесь называется интервью, в ходе которого был получен ответ на последний вопрос, то есть можно предположить, что респондент был готов до конца сотрудничать с исследователем.

Мы провели статистический анализ данных (с помощью коэффициента Хи-квадрат) для каждой волны, чтобы проверить, насколько значимы различия между долями прерванных интервью в случае использования различных приветствий. Результаты анализа представлены в таблице 6. Как видно, лишь в первой и третьей волнах опроса влияние приветствия на долю прерванных интервью статистически значимо, показывая сильную зависимость между переменными. Значение Хи-квадрат: 29,81 в первой волне и 15,23 — в третьей, существенно превышает критическое значение (9,49). Как видим, приветствие № 3, акцентирующее внимание респондента на том, что результаты исследования касаются его собственного благополучия, сильнее прочих вариантов побуждает принимать участие в опросе.

Следует обратить внимание, что, к сожалению, данные, собранные системой и имеющиеся в нашем распоряжении, не позволяют узнать, сколько людей из числа прервавших интервью до первого вопроса успели прослушать приветствие полностью. Поэтому по итогам второй и четвертой волн мы можем судить только о количестве тех респондентов, которые доходили до первого вопроса.

Влияние формулировки вопроса о регистрации избирателя на кооперацию респондента

Второе (и последнее) различие в текстах, зачитываемых роботом респондентам, заключалось в том, что в первом вопросе, устанавливающем, что респондент является жителем Санкт-Петербурга, фигурировали два варианта вопроса: «Зарегистрированы/прописаны (50%/50%) ли Вы в Санкт-Петербурге?» Случаев, где в вопросе звучало слово «зарегистрированы», в нашей выборке 473 (см. табл. 7), из них были прерваны 174 интервью (38%). Случаев, где звучало слово «прописаны», — 373, из них прервано 157 (42%).

Таблица 7. Влияние различных формулировок скрининговых вопросов на кооперацию респондентов

	Прерванное интервью	Завершенное интервью
«Зарегистрированы»	774	258
«Прописаны»	820	197
ВСЕГО	1594	455

Мы наблюдаем наличие статистически значимой зависимости между переменными. Значение Хи-квадрат 21,83 превышает критическое (18,31), а вероятность нулевой гипотезы меньше 5% (1,6%). Как видим, респонденты охотнее продолжают отвечать на вопросы, если слышат слова «зарегистрированы», а не «прописаны».

При этом следует помнить, что формулировка «зарегистрирован» также релевантна случаю регистрации иностранца в местном подразделении миграционной службы. Поэтому данный скринирующий вопрос часто не выполняет свою функцию как для российских граждан, не входящих в целевую группу, так и для иностранных граждан, проживающих в городе.

Основные выводы

Особенности использования роботов в массовых опросах, отмечаемые в работах зарубежных авторов, подтверждаются и на российском материале, представленном в данном исследовании.

Как и в случае проведенных роботизированных опросов в других странах, в России на готовность респондента к кооперации влияет приветствие. При подготовке скрипта можно подобрать слова и формулировки, повышающие вероятность ответа на все вопросы.

Использование роботов позволяет повысить скорость сбора данных за счет того, что можно реализовать практически бесконечное число интервью одновременно, причем в наиболее приемлемый промежуток времени. Наше исследование показало, что для повышения кооперации с респондентами следует выбирать время между окончанием рабочего дня и отходом ко сну. В этот период доля отказов минимальна, а доля контактов с респондентами максимальна.

Большим успехом стало то, что к концу четвертой волны удалось оперативно, за несколько часов, собрать плановую выборку ($N = 1200$). Эти результаты свидетельствуют о том, что роботы оптимально подходят для проведения экспресс-опросов, когда необходимо опросить большое число людей в предельно короткий промежуток времени. Это делает роботов однозначными лидерами не только в скорости опроса, но и в максимальной достижимости целевой аудитории.

Подтвердились и недостатки, присущие IVR. Робот в состоянии поддерживать интерес к беседе лишь в случае, если самому респонденту эта тема интересна. Поэтому очевидно, что дизайн роботизированных интервью не должен предусматривать более 10—12 вопросов, особенно когда зачитывание некоторых из них, включая варианты ответов, занимает много времени, поскольку это может заставить респондента потерять терпение и прервать интервью. Однако наше исследование не подтвердило гипотезу о том, что прерывание интервью вызвано усталостью респондента. Более существенными причинами стали тема интервью и ограниченные возможности по распознаванию речи респондента.

Анализ аудиозаписей интервью показал, что робот далеко не всегда разбирает ответы, особенно когда респондент начинает отвечать до окончания зачитывания вопроса. Робот не готов к нестандартным ответам. В этой связи целесообразно использовать максимально простые и короткие шкалы. При подготовке широкомасштабного интервью крайне желательно провести тест с целью улучшения настроек распознавания ответов и выявления проблемных вариантов. Регулярные опросы с возможностью повторных контактов/звонков (мониторинг) позволяют собрать относительно постоянный объем лексических единиц. Робот плохо распознает ответы на фоне внешних шумов и помех в линиях связи, но эта техническая проблема решается путем улучшения качества программного обеспечения.

Самым главным из полученных нами результатов стало подтверждение того, что роботизированные опросы дают примерно тот же уровень кооперации респондентов и ту же долю прерванных интервью, что и традиционные опросы. Это указывает на то, что, имея большие временные и финансовые преимущества, роботизированные опросы гарантируют качество собираемых данных, и не уступают в этом традиционным формам проведения массовых опросов.

Список литературы (References)

- Рогозин Д. М., Вырская М. С. Коммуникативные сбои и способы их преодоления в роботизированном телефонном опросе // Социологический журнал. 2019. Т. 25. № 2. С. 33—59. <https://doi.org/10.19181/socjour.2019.25.2.6385>.
- Rogozin D. M., Vyrskaya M. S. (2019) Communication Errors and Repairs in the Automated Telephone Interviews. *Sociological Journal*. Vol. 25. No. 2. P. 33–59. <https://doi.org/10.19181/socjour.2019.25.2.6385>. (In Russ.)
- Amaya A., Lau C., Owusu-Amoah Y., Light J. (2018) Evaluation of Gaining Cooperation Methods for IVR Surveys in Low- and Middle-Income Countries. *Survey Methods: Insights From the Field*. <https://doi.org/10.13094/SMIF-2018-00005>.
- Charoenruk N., Olson K. (2018) Do Listeners Perceive Interviewers' Attributes From Their Voices and Do Perceptions Differ by Question Type? *Field Methods*. Vol. 30. No. 4. P. 312–328. <https://doi.org/10.1177/1525822X18784500>.
- Church A. H. (1993) Estimating the Effect of Incentives on Mail Survey Response Rates: A Meta-Analysis. *Public Opinion Quarterly*. Vol. 57. No. 1. P. 62–79. <https://doi.org/10.1086/269355>.
- Comley P. (2006) The Games We Play: A Psychoanalysis of the Relationship Between Panel Owners and Panel Participants. ESOMAR Panels Research Conference 2006, Barcelona, Spain.
- Curtin R., Presser S., Singer E. (2005) Changes in Telephone Survey Nonresponse Over the Past Quarter Century. *Public Opinion Quarterly*. Vol. 69. No. 1. P. 87–98. <https://doi.org/10.1093/poq/nfi002>.
- Dillman D. A., Smyth J. D., Christian L. M. (2014) Internet, Phone, Mail, and Mixed-Mode Surveys: the Tailored Design. Hoboken, NJ: John Wiley & Sons.
- Dillman D. A. (2011) Mail and Internet Surveys: the Tailored Design Method-2007 Update With New Internet, Visual, and Mixed-Mode Guide. Hoboken, NJ: John Wiley & Sons.
- Dillman D. A., Phelps G., Tortora R., Swift K., Kohrell J., Berck J., Messer B. L. (2009) Response Rate and Measurement Differences in Mixed-Mode Surveys Using Mail, Telephone, Interactive Voice Response (IVR) and the Internet. *Social Science Research*. Vol. 38. No. 1. P. 1–18. <https://doi.org/10.1016/j.ssresearch.2008.03.007>.
- Falconi M., Johnston S., Hogg W. (2016) A Scoping Review to Explore the Suitability of Interactive Voice Response to Conduct Automated Performance Measurement of the Patient's Experience in Primary Care. *Primary Health Care Research & Development*. Vol. 17. No. 3. P. 209–225. <https://doi.org/10.1017/s1463423615000407>.
- Groves R. M., Singer E., Corning A. (2000) Leverage-Saliency Theory of Survey Participation: Description and an Illustration. *The Public Opinion Quarterly*. Vol. 64. No. 3. P. 299–308. <https://doi.org/10.1086/317990>.

Groves B. (2004) Interviewer Falsification in Survey Research: Current Best Methods for Prevention, Detection, and Repair of Its Effects. *Survey Research*. Vol. 35. No. 1. P. 1–5. URL: <http://www.srl.uic.edu/Publist/Newsletter/2004/04v35n1.pdf>.

James K. A., Levinson A. H., Li Y., Baumann C., Maffey A. (2018) A Population Sample Registry for Public Health Survey Investigations: Methods, Lessons, and Successes. *International Journal of Social Research Methodology*. Vol. 21. No. 5. P. 565–574. <https://doi.org/10.1080/13645579.2018.1451732>.

Kappelhof J. W. S., De Leeuw E. D. (2019) Estimating the Impact of Measurement Differences Introduced by Efforts To Reach a Balanced Response Among Non-Western Minorities. *Sociological Methods & Research*. Vol. 48. No. 1. P. 116–155. <https://doi.org/10.1177/0049124117701474>.

Kreuter F., Presser S., Tourangeau R. (2008) Social Desirability Bias in CATI, IVR, and Web Surveys: The Effects of Mode and Question Sensitivity. *Public Opinion Quarterly*. Vol. 72. No. 5. P. 847–865. <https://doi.org/10.1093/poq/nfn063>.

Lavrakas P. J. (2008) *Encyclopedia of Survey Research Methods*. Thousand Oaks, CA: Sage Publications. URL: <https://us.sagepub.com/en-us/nam/encyclopedia-of-survey-research-methods/book227714#preview> (accessed: 27.11.2020).

Lee H., Kim S., Couper M. P., Woo Y. (2019) Experimental Comparison of PC Web, Smartphone Web, and Telephone Surveys in the New Technology Era. *Social Science Computer Review*. Vol. 372. No. 2. P. 234–247. <https://doi.org/10.1177/0894439318756867>.

Link S., Dursun A., Karakaya K., Hegelheimer V. (2014) Towards Better ESL Practices for Implementing Automated Writing Evaluation. *CALICO Journal*. Vol. 31. No. 3. <https://doi.org/10.11139/cj.31.3.323-344>.

Lynn P. (2009) Methods for Longitudinal Surveys. In: *Methodology of Longitudinal Surveys*. Chichester: John Wiley & Sons. P. 1–19. <https://doi.org/10.1002/9780470743874.ch1>.

Menold N., Kemper C. J. (2013) How Do Real and Falsified Data Differ? Psychology of Survey Response as a Source of Falsification Indicators in Face-to-Face Surveys. *International Journal of Public Opinion Research*. Vol. 26. No. 1. P. 41–65. <https://doi.org/10.1093/ijpor/edt017>.

Millar M. M., Dillman D. A. (2011) Improving Response to Web and Mixed-Mode Surveys. *Public Opinion Quarterly*. Vol. 75. No. 2. P. 249–269. <https://doi.org/10.1093/poq/nfr003>.

Olson K., Smyth J. D., Wood H. M. (2012) Does Giving People Their Preferred Survey Mode Actually Increase Survey Participation Rates? An Experimental Examination. *Public Opinion Quarterly*. Vol. 76. No. 4. P. 611–635. <https://doi.org/10.1093/poq/nfs024>.

Pariyo G. W., Abigail R. Greenleaf A. R., Gibson D. G., Ali J., Selig H., Labrique A. B., Kibria G. M. A., Khan I. A., Masanja H., Flora M. S., Ahmed S., Hyde A. A. (2019) Does

Mobile Phone Survey Method Matter? Reliability of Computer-Assisted Telephone Interviews and Interactive Voice Response Non-Communicable Diseases Risk Factor Surveys in Low and Middle Income Countries. *PLoS One*. Vol. 14. No. 4. e0214450. <https://doi.org/10.1371/journal.pone.0214450>.

Shaw W. S., Verma S. K. (2007) Data Equivalency of an Interactive Voice Response System for Home Assessment of Back Pain and Function. *Pain Research and Management*. Vol. 12. No. 1. P. 23–30. <https://doi.org/10.1155/2007/185863>.

Singer E., Ye C. (2013) The Use and Effects of Incentives in Surveys. *The ANNALS of the American Academy of Political and Social Science*. Vol. 645. No. 1. P. 112–141. <https://doi.org/10.1177/0002716212458082>.

Singer E. (2011) Toward a Benefit-Cost Theory of Survey Participation: Evidence, Further Tests, and Implications. *Journal of Official Statistics*. Vol. 27. No. 2. P. 379–392. URL: <https://www.scb.se/contentassets/f6bcee6f397c4fd68db6452fc9643e68/toward-a-benefit-cost-theory-of-survey-participation-evidence-further-tests-and-implications.pdf> (accessed: 27.11.2020).

Stafford L. (2008) Social Exchange Theories. In: Baxter L. A., Braithwaite D. O. (eds.) *Engaging Theories in Interpersonal Communication: Multiple Perspectives*. Los Angeles, CA: Sage. P. 377–389.

Trussell N., Lavrakas P. J. (2004) The Influence of Incremental Increases in Token Cash Incentives on Mail Survey Response: Is There an Optimal Amount? *Public Opinion Quarterly*. Vol. 68. No. 3. P. 349–367. <https://doi.org/10.1093/poq/nfh022>.

Willig J. H., Krawitz M., Panjamapirom A., Ray M. N., Nevin C. R., English T. M., Cohen M. P., Berner E. S. (2013) Closing the Feedback Loop: an Interactive Voice Response System to Provide Follow-Up and Feedback in Primary Care Settings. *Journal of Medical Systems*. Vol. 37. No. 2. 9905. <https://doi.org/10.1007/s10916-012-9905-4>.